REMARKS

Claims 1-26 and 29-31 were pending. By the above amendment, the applicants have amended claims 1-4, 8, 10, 15 and 29 and added new claims 33 and 34. The applicants request further consideration and re-examination in view of the amendments above and remarks below.

Rejections under 35 U.S.C. § 112:

Claims 1, 8 and 29 were rejected under 35 U.S.C. § 112, first paragraph, as being indefinite. Specifically, regarding claims 1, 8 and 29, the office action alleges that the specification does not provide a definition for "erasure coded data." The office action further states that the examiner is interpreting this term to mean "any piece of data that can be striped."

The applicants respectfully disagree with the rejection. The use of erasure codes is well known in the art. For example, an Internet search for the terms "erasure codes" and "erasure coded" each turns up numerous results. Reed-Solomon is a well-known example of an erasure code. Accordingly, persons of ordinary skill in the art will certainly understand what is meant by "erasure coded data." Particularly, an erasure code transforms a message of m data blocks into a message having m data blocks and p parity blocks. If one of the data blocks is lost, it can be retrieved by decoding the lost data from a parity block and the remaining data blocks. See applicants' specification at page 2, lines 1-6.

Therefore, the applicants respectfully disagree that the claims are indefinite for inclusion of the term "erasure coded." The applicants also respectfully disagree with the examiner's interpretation of the term erasure coded data as meaning "any piece of data that can be striped." While it is true that erasure coded data can be striped, not all striped data is erasure coded.

In addition, regarding claims 1 and 29, the office action alleges that the specification "does not provide proper support or an explanation as to how the quorum meets a quorum condition and how any two selections of the number of stripe blocks intersect in the minimum number." The office action further states that the limitation is being interpreted by the examiner as "the quorum of storage devices comprising of the minimum number of storage devices necessary for the recovery of the data; wherein the quorum condition is met when the minimum number of storage

devices necessary are available for use in the decoding of the stripe of erasure coded data."

The applicants respectfully disagree with the rejection and with the examiner's interpretation of the claim language. The applicants' specification clearly explains that data is stored across n of the storage devices as a stripe of erasure coded data comprising n stripe blocks. The n stripe blocks may comprise m data blocks plus p parity blocks ($n = m + p$). The p parity blocks may comprise 2f parity blocks, where f is a number of the n stripe blocks that are tolerated as faulty ($n = m + p = m + 2f$). A quorum is defined as a number of the n stripe blocks that meets a quorum condition. The applicants define their quorum condition such that any selection of two sets of the n stripe blocks intersect in m of the n stripe blocks. Therefore, if the n stripe blocks are encoded as the m data blocks plus the p parity blocks and p is an even number, the quorum is $m + p/2$. If the n stripe blocks are encoded as the m data blocks plus the p parity blocks and p is an odd number, the quorum is $m + p/2 + 1/2$. See applicants' specification at page 4, line 22 to page 5, line 2.

It is clear from this detailed explanation contained in the applicants' specification that that applicants have defined the quorum condition to be a number that is greater than m (the number of data blocks in the stripe, excluding parity blocks) and less than n (the total number of blocks in the stripe including data blocks and parity blocks). Specifically, the quorum condition is defined in the specification and in claims 1 and 29 to be a number such that any two selections of the number of the stripe blocks intersect in a minimum number of the stripe blocks needed to decode the stripe. Accordingly, the write transaction of claims 1 and 29 requires a minimum number of participants to reply to the queries specified in the claims. However, the write transaction can still be performed even if a reply is not received from every storage device for the stripe.

To clarify this concept even further, the applicants provide the following example. Suppose that the number of data blocks, m, is four (i.e. $m = 4$) and the number of parity blocks, p, is three (i.e. $p = 3$). Therefore, in this example, the set of stripe blocks has seven members ($n = 7$, since $n = m + p$). Because p is an odd number, the quorum is six (i.e. $m + p/2 + 1/2 = 6$). It can be verified that six is the correct quorum since a first selection of six of the seven blocks leaves one block unselected. A second, different selection of six of the seven blocks would leave one other block unselected. Therefore, the two selections of six of the seven blocks

intersect in five of the blocks. In other words, five blocks must be included in both selections of six of seven blocks. These five blocks are sufficient to decode the stripe. In fact, in this example, four blocks would be sufficient to decode the stripe. Therefore, in this example, the quorum is greater than the minimum number of blocks needed to decode the stripe. Conversely, it can be verified that a number less than six, such as five, does not meet the quorum condition. This is because a first selection of five of the seven blocks would leave two blocks unselected and a second selection of five of the seven blocks could leave as many as two different blocks unselected. Accordingly, the intersection of these two selections would be only three blocks. In other words, only three blocks must be included in both selections of five of seven blocks. However, three blocks is not sufficient to decode the stripe (since the stripe has m=4 data blocks). Therefore, five does not meet the quorum condition. Any number less than five would also not meet the quorum condition.

In view of the above, the applicants respectfully submit that the quorum condition specified in claims 1 and 29 is not indefinite.

Regarding claim 8, the office action alleges that the specification "does not provide proper support or an explanation as to how the quorum can comprise of a first number of data blocks and **half** of the parity blocks." The office action further states that this limitation is being interpreted by the examiner as "the quorum comprising of some of the data blocks plus the parity blocks."

The applicants respectfully disagree with the rejection and with the interpretation of the phrase by the examiner. As explained in the applicants' specification at page 5, lines 9-15, the number of data blocks is given as m. If the number of parity blocks, p, is even, the quorum is m + p/2 and, if p is odd, the quorum is m + p/2 + ½. Therefore, in both cases (where p is even or odd) the quorum comprises *at least* the first number of data blocks m and half of the parity blocks p. Therefore, the limitation of claim 8 that the quorum comprises the first number of data blocks plus a half of the second number of parity blocks is fully supported by the applicants' specification. To further clarify this phrase in claim 8, the applicants have amended claim 8 so that the term "at least" precedes the phrase.

The applicants remind the examiner that according to the Manual of Patent Examining Procedure (MPEP), the essential inquiry regarding claim definiteness is "whether the claims set out and circumscribe a particular subject matter with a reasonable degree of clarity and particularity." Definiteness of claim language must

be analyzed, not in a vacuum, but in light of: the content of the particular application disclosure; the teachings of the prior art; and the claim interpretation that would be given by one possessing the ordinary level of skill in the pertinent art at the time the invention was made. Some latitude in the manner of expression and the aptness of terms should be permitted even though the claim language is not as precise as the examiner might desire. Examiners are encouraged to suggest claim language to applicants to improve the clarity or precision of the language used, but should not reject claims or insist on their own preferences if other modes of expression selected by applicants satisfy the statutory requirement. MPEP at Section 2173.02 (Aug. 2006).

In view of the above, the applicants request that the rejection of claims 1, 8 and 29, under 35 U.S.C. 112 be removed.

Rejections under 35 U.S.C. § 102:

Claims 1, 4-7 and 29 were rejected as being anticipated by U.S. Patent Publication No. 2002/0083379 by Nishikawa et al (hereinafter "Nishikawa"). Particularly, claims 1 and 29 are rejected in reliance on paragraphs 7-9 and 113 of Nishikawa.

As discussed above, the quorum recited in claims 1 and 29 comprises at least the minimum number of data blocks required to decode the stripe and can comprise a number that is greater than the minimum number of blocks needed to decode the stripe. To further clarify this point, the applicants have amended claims 1 and 29 to specify that the quorum of the reply messages is greater than the minimum number of stripe blocks needed to decode the data. This is useful because if only the minimum number of replies are received that are needed to decode the data, this means that a failure of one of those devices could result in a situation where the data cannot be decoded. Therefore, as explained in the applicants' specification at page 7, lines 14-20, if at least a quorum of replies are received, the data is decoded. And, as explained in the applicants' specification at page 7, lines 21-26, if less than the quorum of replies are received, a data recovery operation is performed. From this, it is clear that the data is decoded in response to receiving the quorum of replies and that not all of the devices need to respond in order for the data to be decoded. Accordingly, the applicants have also amended claims 1 and 29 to clarify that the data is decoded in

response to receiving at least the quorum of reply messages and not necessarily a reply from each of the storage devices of the plurality.

In paragraphs 7-9, and 113, Nishikawa discloses a conventional RAID 3 data storage server system. In accordance with this scheme, data units are stored in each of four storage units and a fifth storage unit stores parity data. When data is written to the system, the data is divided and stored in each of the four storage units and the parity data is generated from the data and stored in the fifth storage unit. When the data is read, the data and parity data is retrieved from all of the storage units that store the data. If one of the storage units that stores data has failed such that its portion of the data is lost, the lost data is recovered from the data stored by remaining three storage units and from the parity data.

Applicants' claims 1 and 29 recite features which are not disclosed by Nishikawa. For example, Nishikawa teaches in paragraph 7 that in response to receiving a read request, data from all five data storage units is read out and transferred to the host; if one of the data storage units has failed, the lost portion is recovered from the remaining four data storage units. Therefore, Nishikawa does not teach that a "quorum" of replies (which is greater than the minimum number needed to decode the data) is required to be received before the data is decoded. Instead, Nishikawa teach decoding the data when replies are received from only four of the storage devices. However, four does not meet the quorum condition specified in applicants' claim 1 and 29 since two selections of four of the five data units would intersect in only three data units, whereas, four is the minimum number needed to decode the data.

For at least this reason, applicants' claims 1 and 29 are allowable over Nishikawa. Claims 4-7 are allowable at least because each depends from an allowable base claim 1.

Dependent claim 4 is amended to recite that the storage devices are distributed with each storage device comprising a CPU and storage and wherein a coordinator performs the steps. The amendments to claim 4 are supported by the applicants' specification at least at page 4, lines 19-27. In contrast to claim 4, Nishikawa teaches a centralized server containing a single controller. Therefore, Nishikawa does not teach that the storage devices are distributed, nor does not Nishikawa teach that each storage device comprises a CPU and storage. This is another reason why claim 4 is allowable is also another reason why claims 5-7 are allowable, being dependent from

claim 4. Moreover, the reasons given in the office action as to why claims 5-7 are allegedly anticipated do not apply considering that the storage devices are distributed.

Rejections under 35 U.S.C. § 103:

Claims 2, 8, 10-15, 17-23, 25-26 and 30 were rejected as being unpatentable over Nishikawa in view of U.S. Patent Publication No. 2004/0158677 by Dodd (hereinafter "Dodd").

Regarding claim 2, the office action states that Nishikawa does not teach each of the reply messages within the quorum indicating that there is no pending write for the stripe block stored on the storage device associated with the reply message. However, the office action alleges that Dodd teaches this feature at paragraph 18 and that it would have been obvious to incorporate this feature of Dodd "thus avoiding the reading of data that is not up to date (data that needs to be modified by a pending write)."

The applicants respectfully traverse the rejection. Claim 2 is dependent from an allowable base claim 1. For at least this reason claim 2 is allowable. Further, as explained below, the applicants respectfully submit that claim 2 recites limitations that are not obvious in view of the cited art.

Dodd explains in paragraph 1 that performance of a computing device is strongly influenced by the memory latency of the device. As is also explained by Dodd, "memory read latency" is the length of time between when the processor requests the memory controller to retrieve data from memory and when the memory controller provides the data to the processor. "Memory write latency" is generally the length of time between when the processor requests the memory controller to write data to the memory and when the controller indicates to the processor that the data has been or will be written to the memory. In paragraph 2, Dodd explains that once the write transaction is buffered by the memory, the processor may continue without waiting for the data of the write transaction to be actually written to memory. Conversely, for a read transaction, the processor must wait until the data is read from memory and returned to the processor. Therefore, performance of a computing device is typically more dependent upon read latency than write latency. Dodd explains that in light of this, memory controllers tend to favor servicing read transactions over servicing write transactions.

In paragraph 17, Dodd explains that a read buffer may store the address and

data for read transactions until the requested data is retrieved and returned to the requestor. Similarly, a write buffer may store the address and data for write transactions until the data is written to the memory. Because read transactions are favored over write transactions, the write transactions generally have to wait until the read buffer is empty. However, as explained in paragraph 18 of Dodd, write transactions may be selected over read transactions under certain conditions, such as the write buffer becoming full or nearly full.

Therefore, Dodd teaches that buffered read transactions are generally favored over buffered write transactions, but that write transactions may be performed when there are still read transactions in the read buffer if the write buffer is full or nearly full. However, Dodd teaches nothing about individual stripe blocks of erasure coded data. As such, Dodd cannot teach anything about reply messages associated with individual stripe blocks of erasure coded data. However, these features are recited in applicants' claim 2.

Further, claim 2 is amended to recite that each of the reply messages is included within the quorum only if it indicates that there is no pending write for the stripe block stored on the storage device associated with the reply message. These amendments are supported by the applicants' specification at page 7, lines 14-20. Dodd also does not teach anything about the quorum specified in the applicants' claims, nor does Dodd teach anything about conditions under which reply messages are included within such a quorum.

Nishikawa does not disclose the limitations that are missing from Dodd. Therefore, applicants' claim 2 recites limitations not found in Nishikawa or Dodd, taken singly or in combination. The applicants respectfully submit that these are additional reasons that claim 2 is allowable.

Further, the applicants respectfully submit that it would not have been obvious to combine Nishikawa with Dodd in the manner suggested in the office action. For example, the office action alleges that it would have been obvious to combine the references "thus avoiding the reading of data that is not up to date (data that needs to be modified by a pending write)." However, this alleged motivation is unrelated to the reasons why Dodd might chose to perform a write transaction while there are still read transactions in the read buffer. Instead, Dodd teaches that this choice is based on the write buffer being full or nearly full. It appears that the alleged motivation does not come from the prior art at all, but instead comes from the applicants' own

disclosure. This is because it is the applicants, not the cited art, who state that checking that there is no pending write for the stripe blocks is done to ensure data integrity. See applicants' specification at page 16, lines 29-32. As stated in the Manual of Patent Examining Procedure, the applicant's own disclosure cannot form the basis for suggesting a particular combination of references. See MPEP at Section 2143 (Aug., 2006).

Independent claim 8 is rejected using essentially the same rationale as claim 2. Particularly, the office action asserts that Nishikawa teaches all of the limitations of claim 8 except that Nishikawa does not teach each of the reply messages within the quorum indicating that there is no pending write for the stripe block stored on the storage device associated with the reply message. However, the office action asserts that Dodd discloses this feature and that it would have been obvious to combine Nishikawa with Dodd for the same reason given with respect to claim 2.

The applicants respectfully traverse the rejection. First, Nishikawa does not disclose "receiving at least a quorum of reply messages from the storage devices ... the quorum comprising at least the first number plus a half of the second number and the quorum being greater than a minimum number of the stripe blocks needed to decode the stripe of erasure coded data," as recited in applicants' claim 8. Claim 8 is amended similarly to claim 1 to recite that the quorum is greater than a minimum number of the stripe blocks needed to decode the stripe of erasure coded data. However, Nishikawa does not disclose such a quorum because Nishikawa simply reads out the data from the four storage devices that store the requested data. And, in the arrangement of Niskikawa, four is the minimum number needed to decode the data.

Dodd does not disclose this feature of claim 8 that is missing from Nishikawa. For at least this reason, claim 8 is allowable over Nishikawa and Dodd, taken singly or in combination.

Further, claim 8 recites that the reply messages from the storage devices indicate that there is no pending write for the stripe block stored on the storage device. However, as explained above in connection with claim 2, Dodd teaches nothing about individual stripe blocks of erasure coded data. As such, Dodd cannot teach anything about reply messages associated with individual stripe blocks of erasure coded data. Dodd also does not teach anything about the quorum specified in the applicants' claims, nor does Dodd teach anything about conditions under which reply messages

are included within such a quorum.

Nishikawa does not disclose the limitations that are missing from Dodd. Accordingly, the applicants submit that these are additional reasons that claim 8 is allowable over Nishikawa and Dodd.

Further, the applicants submit that it would not have been obvious to combine Nishikawa with Dodd in the manner suggested in the office action. This is because the motivation to make the combination that is alleged in the office action (i.e. avoiding the reading of data that is not up to date) is unrelated to the reasons why Dodd might chose to perform a write transaction while there are still read transactions in the read buffer. Instead, Dodd teaches that this choice is based on the write buffer being full or nearly full.

This is another reason why claim 8 is allowable.

Claims 10-15, 17-23 and 25-26 are allowable at least because they depend from an allowable base claim 8. Further, these dependent claims recite limitations which are not obvious in view of the cited art. For example, claim 10 is amended to recite that the storage devices are distributed with each storage device comprising a CPU and storage and wherein a coordinator performs the steps. The amendments to claim 10 are supported by the applicants' specification at least at page 4, lines 19-27. Claim 19 also recites that the storage devices comprise a distributed storage system. In contrast to claims 10 and 19, Nishikawa teaches a centralized server containing a single controller. This is another reason why claims 10 and 19 are allowable and is also another reason why claims 11-14 and 20-23 are allowable, being dependent from claim 10.

Claims 3 and 31 were rejected as being unpatentable over Nishikawa in view of U.S. Patent No. 6,973,556 to Milligan et al. (hereinafter "Milligan"). Claims 9 and 24 were rejected as being unpatentable over Nishikawa in view of Dodd and further in view of Milligan. Claim 16 was rejected as being unpatentable over Nishikawa in view of Dodd and further in view of U.S. Patent Publication No. 2004/0064633 by Oota (hereinafter "Oota").

Claims 3, 9, 16, 24 and 31 are allowable at least because each depends from an allowable base claim. Further, these dependent claims recite limitations which are not obvious in view of the cited art. For example, claims 3, 9, 24 and 31 recite the use of specified timestamps in specified manners. Specifically, claim 3 recites that "each of the reply messages is included within the quorum only if it indicates that the stripe

block associated with the reply message has a timestamp that matches other timestamps associated with other reply messages within the quorum." Claim 9 recites that "the quorum of the reply messages includes validation timestamps which match." Claim 24 recites that "each of the storage devices comprises a log, wherein the log comprises log entries of each successful write of data, the log entries comprising a stripe indicator, a write timestamp, and a physical location of the stripe block on the storage device." Claim 31 recites that "each of the reply messages within the quorum indicate that the stripe block associated with the reply message has a timestamp that matches other timestamps associated with other reply messages within the quorum." Milligan discloses timestamps indicating the time or relative sequence number "of some event or events associated with this data element." See Milligan at column 5, line 61 to col. 6, line 4. However, this does not teach or suggest the specific limitations of applicants' claims 3, 9, 24, and 31, nor does the office action attempt to show how the specific limitations of claims 3, 9, 14 and 31 are alleged to be taught by Milligan. Therefore, these are additional reasons why claims 3, 9, 14 and 31 are allowable.

New Claims:

New claims 33 and 34 are dependent from claims 1 and 8, respectively. Claims 33 and 34 recite that in response to receiving fewer than the quorum of replies, the stripe of erasure coded data is recovered by writing the stripe of erasure coded data to a number of the storage devices that at least meets the quorum condition. This limitation is supported by the applicant's specification at least at page 7, lines 21-26, and page 10, line 14, to page 13, line 3. New claims 33 and 34 are allowable at least because each depends from an allowable base claim.

Conclusion:

In view of the above, the Applicants submit that all of the pending claims are now allowable. Allowance at an early date would be greatly appreciated. Should any outstanding issues remain, the Examiner is encouraged to contact the undersigned at

(408) 293-9000 so that any such issues can be expeditiously resolved.

Respectfully Submitted,

Dated: June 12, 2007

Derek J. Westberg (Reg. No. 40,872)